.

SMICA

J.-F. Cardoso, C.N.R.S.

Institut d'Astrophysique de Paris

Tenerife, Oct. 218

Not SMICA

Saving My Incredibly Cool Asteroid

SMICA = SM + ICA

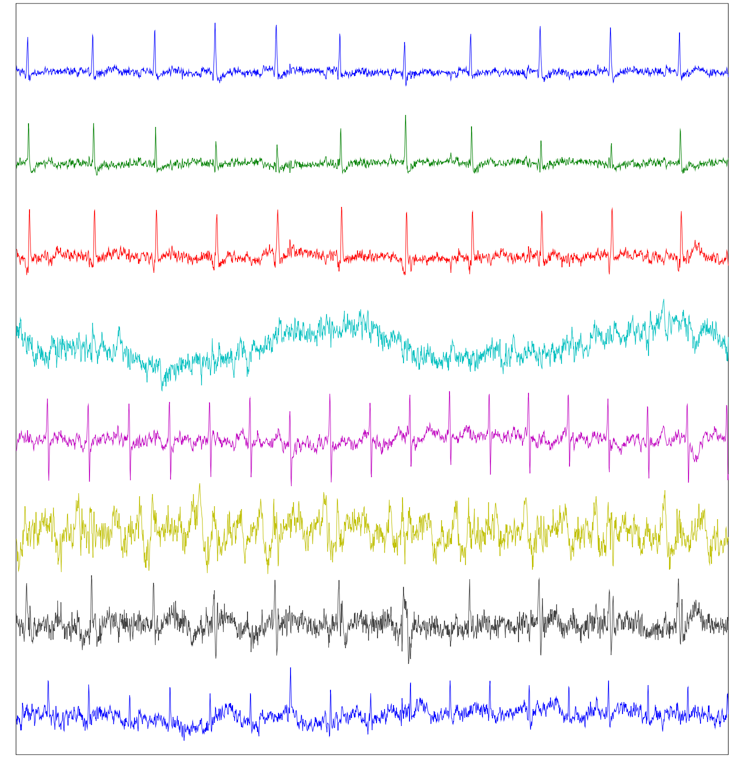
SM = Spectral Matching

ICA = Independent Component Analysis

Independent Component Analysis

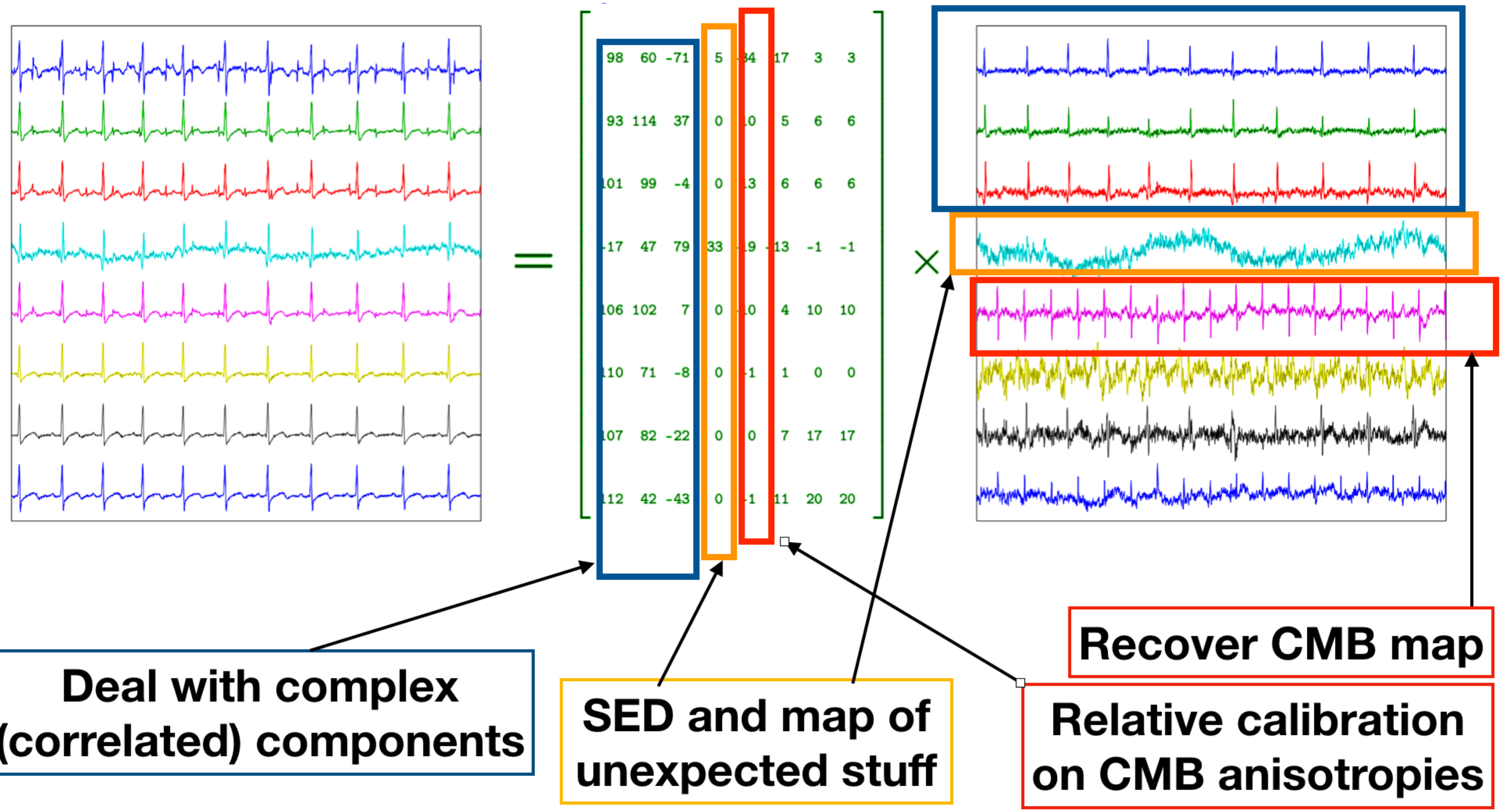
Independent Component Analysis (here: toy model on real ECG data)

$$= \begin{bmatrix} 98 & 60 & -71 & 5 & -34 & 17 & 3 & 3 \\ 93 & 114 & 37 & 0 & 10 & 5 & 6 & 6 \\ 101 & 99 & -4 & 0 & 13 & 6 & 6 & 6 \\ -17 & 47 & 79 & 33 & -19 & -13 & -1 & -1 \\ 106 & 102 & 7 & 0 & -10 & 4 & 10 & 10 \\ 110 & 71 & -8 & 0 & -1 & 1 & 0 & 0 \\ 107 & 82 & -22 & 0 & 0 & 7 & 17 & 17 \\ 112 & 42 & -43 & 0 & -1 & 11 & 20 & 20 \end{bmatrix} \times$$



- Linear decomposition into "the most independent sources"
- Blind: only independence is at work but it must go beyond decorrelation.
- Independence is statistically very strong but often physically plausible.
- Weak assumptions $\longrightarrow$ wide applicability.

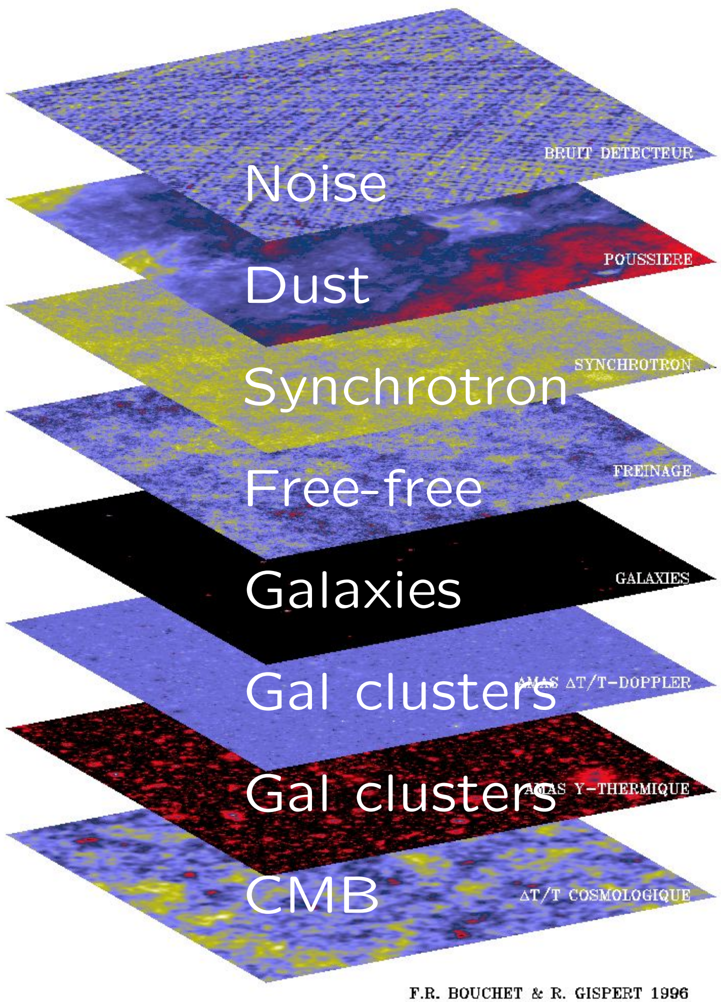
ICA for radio-astronomy



- Good enough for ‘precision cosmology’ ?

$\rightarrow$ Let’s do it well, with a better model and a decent likelihood.

Good old noisy linear model



Foreground emissions pile up in front of the CMB but they do so additively (no occlusion) and most scale rigidly with frequency. Hence, the model in the harmonic domain, after beam correction:

$$\mathbf{d}_{\ell m} = \mathbf{A}\mathbf{s}_{\ell m} + \mathbf{n}_{\ell m} \qquad \mathbf{d}_{\ell m} : N_{\mathsf{chan}} \times 1$$

Mixing matrix $\mathbf{A}$ is $N_{\mathsf{chan}} \times N_{\mathsf{comp}}$.
Each column of $\mathbf{A}$ is a SED (**or not!**).
Add columns as needed for partially coherent emissions or spatially varying SEDs.

Such a linear mixture can be inverted . . . if the mixing matrix $\mathbf{A}$ is known. How to find it or do without it ?
1 Trust astrophysics and use parametric models, or
2 Trust your data and the power of statistics.

Example: the SMICA model for the 2013 CMB temperature map

- Use all 9 Planck channels and capture foreground emission with 6 "foregrounds":

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_9 \end{bmatrix} = \begin{bmatrix} a_1 & F_{11} & \dots & F_{16} \\ a_2 & F_{21} & \dots & F_{26} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ a_9 & F_{91} & \dots & F_{96} \end{bmatrix} \times \begin{bmatrix} s \\ f_1 \\ \vdots \\ f_6 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ \vdots \\ n_9 \end{bmatrix} \quad \text{that is} \quad \mathbf{d}_{\ell m} = [\,\mathbf{a} \mid \mathbf{F}\,] \begin{bmatrix} s_{\ell m} \\ \mathbf{f}_{\ell m} \end{bmatrix} + \mathbf{n}_{\ell m}$$

- SMICA only assumes <u>decorrelation</u> between foregrounds and CMB:

$$\mathrm{Cov}\left(\begin{bmatrix} s_{\ell m} \\ \mathbf{f}_{\ell m} \end{bmatrix} \right) = \begin{bmatrix} C_\ell^{\mathrm{cmb}} & 0 \\ 0 & \mathbf{P}_\ell \end{bmatrix}.$$

The foregrounds must have 6 (say) dimensions but are otherwise completely unconstrained: they may have any spectrum, any color, any correlation. . .

So the data model is **very blind**: all non-zero parameters* are <u>free</u> !

$$\mathrm{Cov}(\mathbf{d}_{\ell m}) = [\,\mathbf{a} \mid \mathbf{F}\,] \begin{bmatrix} C_\ell^{\mathrm{cmb}} & 0 \\ 0 & \mathbf{P}_\ell \end{bmatrix} [\,\mathbf{a} \mid \mathbf{F}\,]^\dagger + \begin{bmatrix} \sigma_{1\ell}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{9\ell}^2 \end{bmatrix}.$$

One subspace to rule them all (out).

*

Spectral matching

Spectral matching in a nutshell

1 From N input maps, compute spectral $N \times N$ covariance matrices

$$\widehat{\mathbf{C}}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{m=+\ell} \mathbf{d}_{\ell m} \mathbf{d}_{\ell m}^\dagger$$ contains all empirical auto- and cross-spectra.

2 Think of a model for them: $\theta \to \mathbf{C}_\ell(\theta) \stackrel{\text{def}}{=} \mathrm{Cov}_\theta(\mathbf{d}_{\ell m}) = \mathbb{E}_\theta \widehat{\mathbf{C}}_\ell$.

3 Fit the model *i.e.* minimize a spectral mismatch between data and model:

$$\widehat{\theta} = \arg\min_\theta \sum_\ell (2\ell+1)\ \mathbf{K}\left[\widehat{\mathbf{C}}_\ell,\ \mathbf{C}_\ell(\theta)\right]$$

4 Dissect the fitted model $\mathbf{C}_\ell(\widehat{\theta})$ and/or build component maps.

FAQ: OK, you jointly fit all auto- and cross-spectra, but...
. Why $\mathbf{K}[\cdot,\cdot]$ rather than a squared difference $\widehat{\mathbf{C}}_\ell - \mathbf{C}_\ell(\theta)$?
. And where is your error covariance matrix ?
. And where is your χ^2 ?

.

Special K

The probability p of observing a given realization of N jointly Gaussian stationary fields on the sphere with spectral covariance matrices $\mathbf{C}_\ell$ is given by

$$-\log p = \sum_\ell (2\ell+1)\,\mathbf{K}\left[\widehat{\mathbf{C}}_\ell, \mathbf{C}_\ell\right] + \mathrm{cst}$$

with $\widehat{\mathbf{C}}_\ell = \frac{1}{2\ell+1}\sum_m \mathbf{d}_{\ell m}\mathbf{d}_{\ell m}^\dagger$ and where for any 2 positive matrices A and B:

$$\mathbf{K}[A,B] \stackrel{\text{def}}{=} [\mathrm{trace}(AB^{-1}) - \log\det(AB^{-1}) - N]/2.$$

$\rightarrow$ SMICA's spectral mismatch is just a Gaussian (in the maps) likelihood.

For small Gaussian (in the spectra) errors,

$$(2\ell+1)\,\mathbf{K}\left[\widehat{\mathbf{C}}_\ell, \mathbf{C}_\ell\right] \approx \frac{1}{2}\,\mathrm{Vec}\left(\widehat{\mathbf{C}}_\ell - \mathbf{C}_\ell\right)^\dagger\ \mathrm{Cov}_G(\mathrm{Vec}(\widehat{\mathbf{C}}_\ell))^{-1}\ \mathrm{Vec}\left(\widehat{\mathbf{C}}_\ell - \mathbf{C}_\ell\right).$$

$\rightarrow$ SMICA's has its own built-in implicit optimal (???) error weighting machine.

.

SMICA in polarization

Use T modes

Or use E modes

Or B modes

Or B+E modes

Or T+E+B modes

No specific differences in principle.

Making component maps

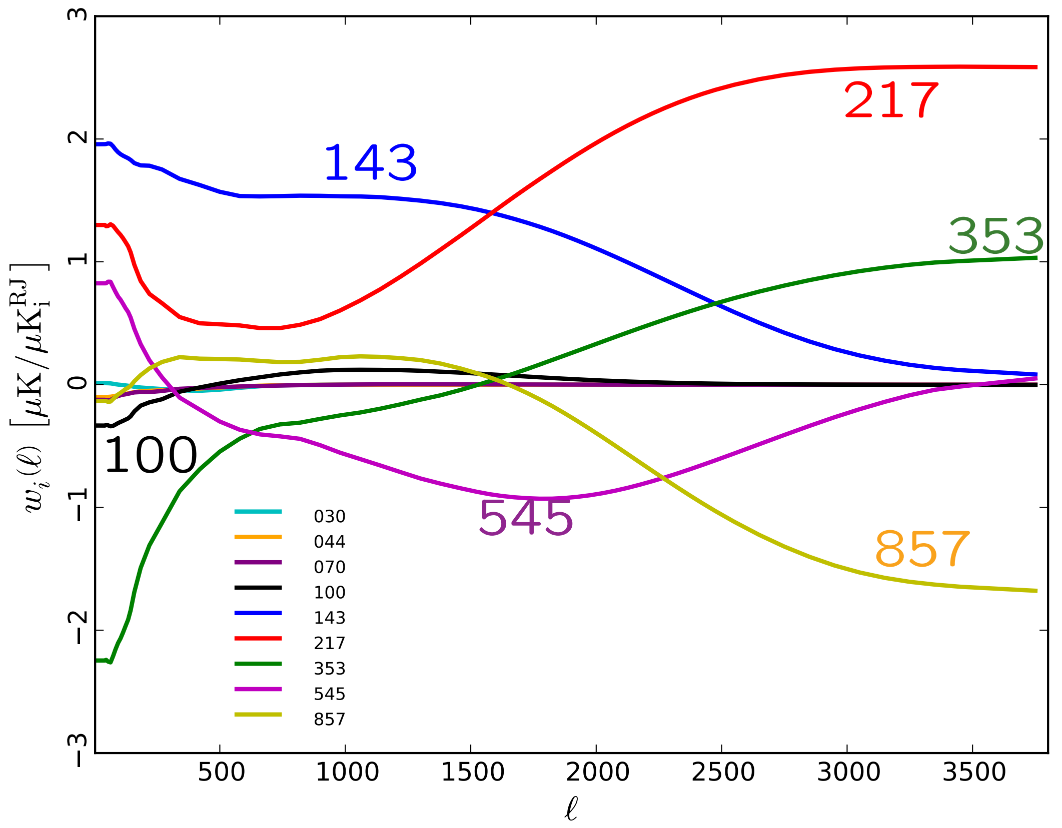
Linear filtering in harmonic space

Since resolution, noise and foregrounds vary (wildly) in power over channels and angular frequency, the combining weights should depend on ℓ.

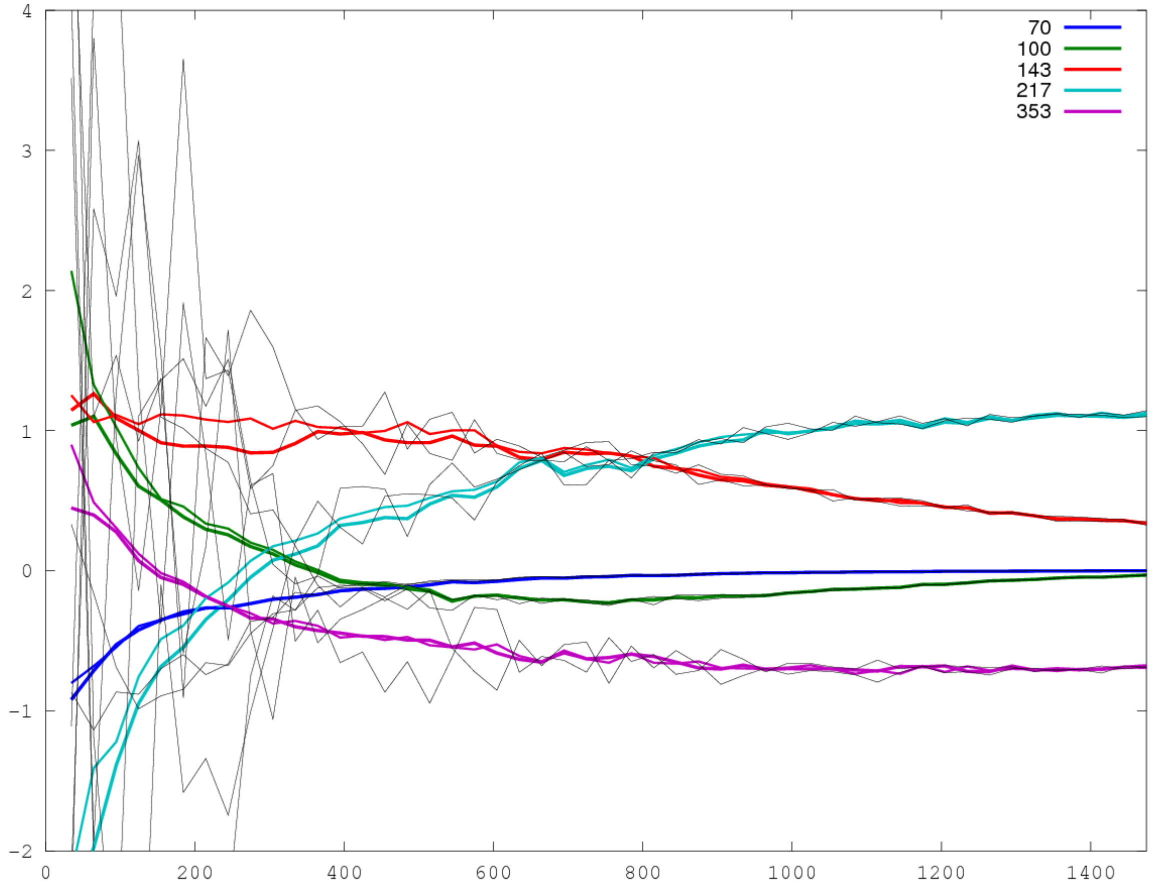
The 'optimal' harmonic combination of maps would be synthesized from $\widehat{s}_{\ell m} = \mathbf{w}_\ell^\dagger \, \mathbf{d}_{\ell m}$ with weights

$$\mathbf{w}_\ell = \frac{\mathbf{C}_\ell^{-1}\,\mathbf{a}}{\mathbf{a}^\dagger \mathbf{C}_\ell^{-1}\,\mathbf{a}} \qquad \mathbf{C}_\ell = \mathrm{Cov}(\mathbf{d}_{\ell m})$$

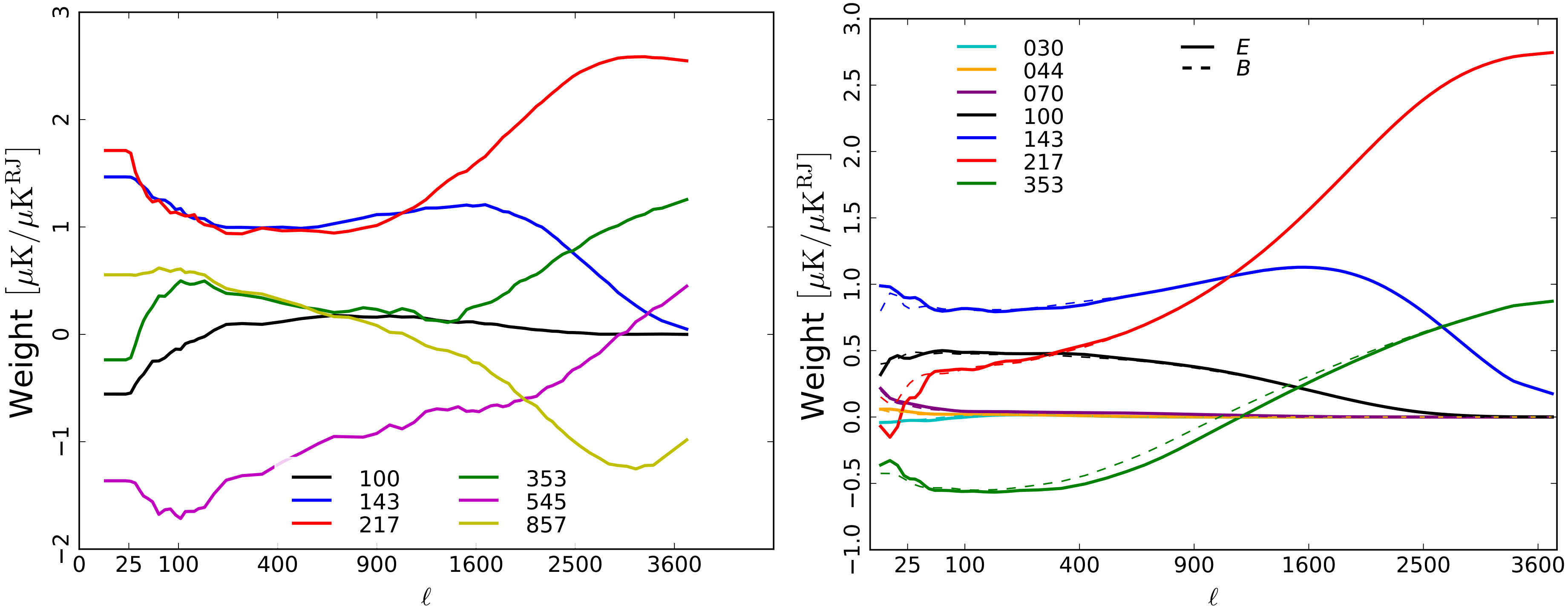
Instead of $\mathbf{C}_\ell$, SMICA maps are build using $\mathbf{C}_\ell(\widehat{\theta})$, as fitted by SMICA.



ILC coefficients: raw (thin lines) and via SMICA modelling (thick lines)



Spectral weights for CMB extraction in the 2018 release



Left: Temperature.

Right: Polarization. Weights on E and B modes.

Separating polarized foregrounds, almost blindly.

Dedicated SMICA fits on E and B spectra with a 2-dimensional foreground model:

$$\mathrm{Cov}(\mathbf{d}_{\ell m}) = [\,\mathbf{a} \mid \mathbf{F}\,] \begin{bmatrix} C_\ell^{\mathrm{cmb}} & 0 \\ 0 & \mathbf{P}_\ell \end{bmatrix} [\,\mathbf{a} \mid \mathbf{F}\,]^\dagger + \begin{bmatrix} \sigma_{1\ell}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{9\ell}^2 \end{bmatrix}$$

The foreground contribution is $\mathbf{F}\mathbf{P}_\ell\mathbf{F}^\dagger$ with $\mathbf{F} : 7 \times 2$ and $0 \leq \mathbf{P}_\ell : 2 \times 2$. Note that

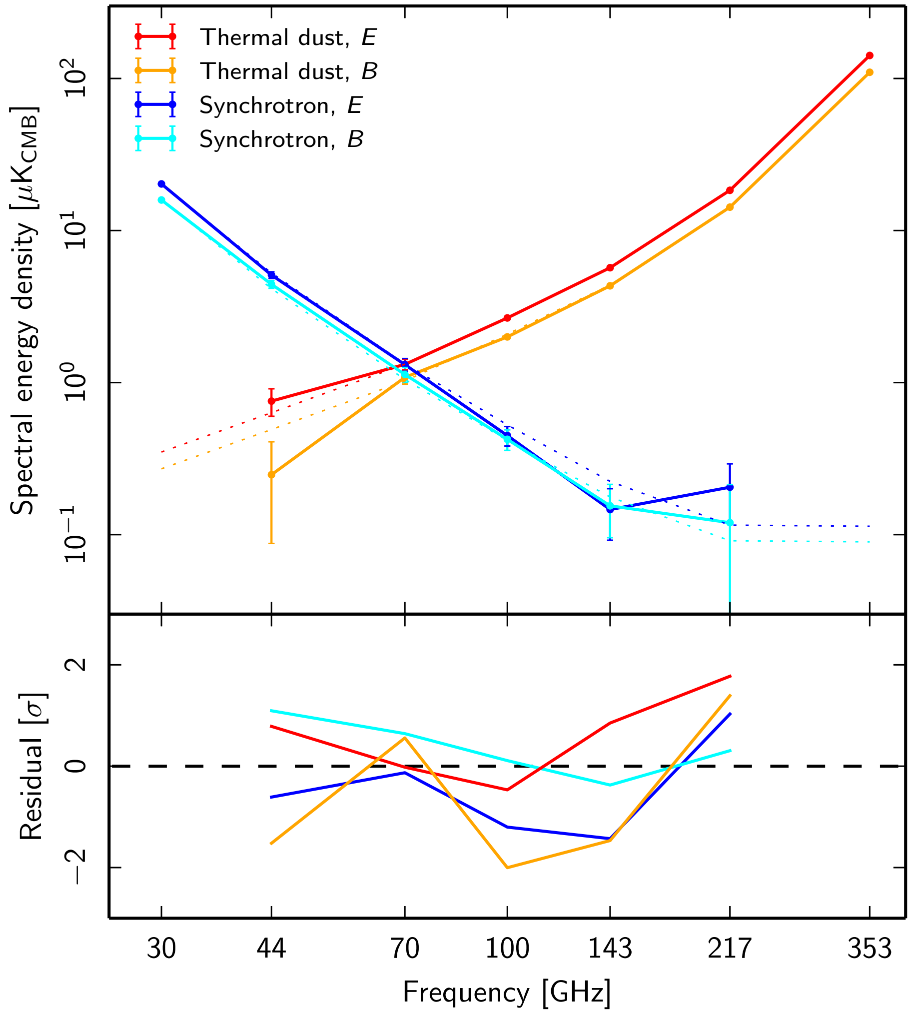
$$\mathbf{F}\mathbf{P}_\ell\mathbf{F}^\dagger = \mathbf{F}\ \mathbf{T}\ \mathbf{T}^{-1}\ \mathbf{P}_\ell\ \mathbf{T}^{-\dagger}\ \mathbf{T}^\dagger\ \mathbf{F}^\dagger \quad \text{for any invertible } 2 \times 2 \text{ matrix } \mathbf{T}.$$

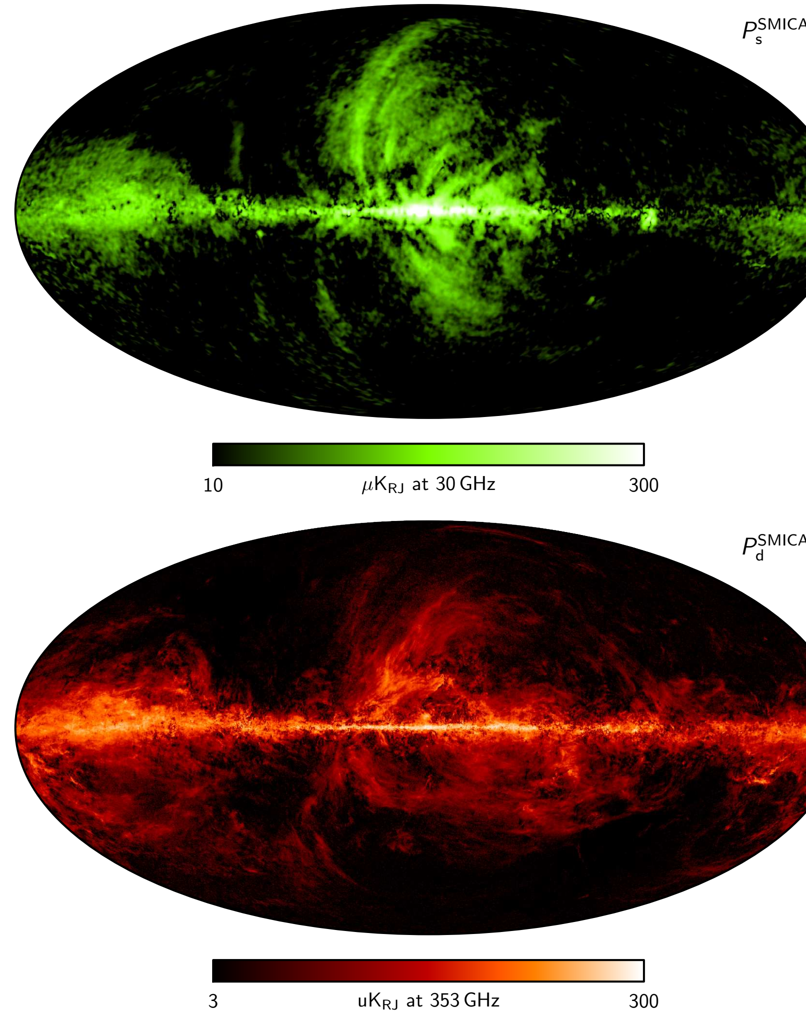
Meaning: synchrotron and dust cannot be disentangled blindly because we cannot/do not want to impose a diagonal matrix $\mathbf{P}_\ell$ (correlation).

But we can choose $\mathbf{T}$ so that one column of $\mathbf{FT}$ has a zero emission at 30 GHz and the other at 353 GHz.

In this manner, we 'almost blindly' separate synchroton and dust.

Separating polarized foregrounds, *post hoc*





Best *post hoc* fits: $\beta_s = -3.10 \pm 0.06$ and $\beta_d = 1.53 \pm 0.01$, $T_d = 19.6K$

Goodies

Goodies 1) Calibration and mis-calibration

Early planet-based miscalibration in Planck detected by SMICA.

Now routinely running SMICA in ‘relative calibration mode’:
estimate spectral matrices on relatively clean fractions of the sky
and let SMICA freely adjust the calibration coefficients (*i.e.* vector $\mathbf{a}$ above)
on the first CMB peaks.

Relative accuracy of a few per mille on cosmological channels.

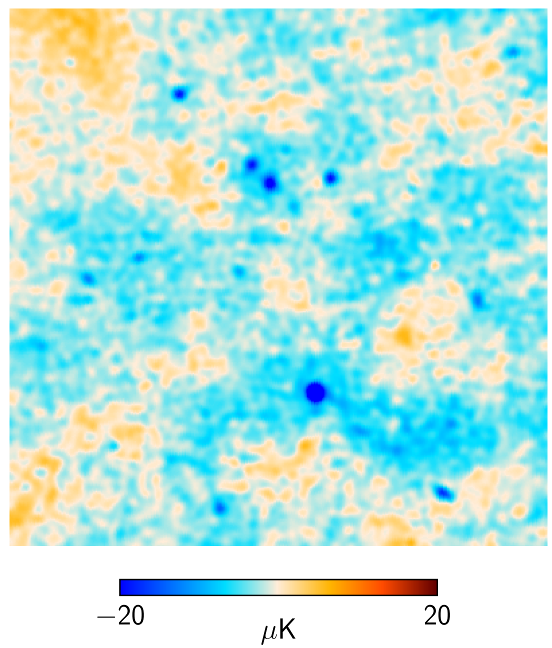
Used to check consistency with calibration based on solar dipole

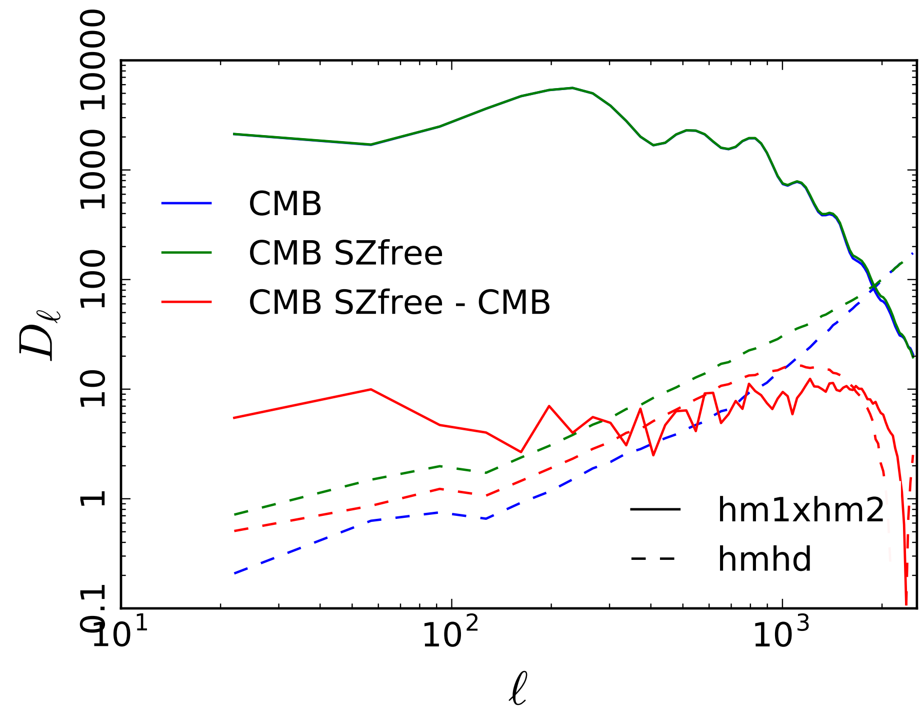
Same procedure for determination of polar efficiency.

See Planck 2018 results. III. HFI data processing and frequency maps.

Goodies 2) Nulling out components with known emissivity

Harmonic ILC trivially adapted to get minimum variance unbiased CMB and additionally null out a component with known SED.





Difference between the SMICA CMB map and its SZ-free version. Patch is $20^\circ \times 20^\circ$ centered on $(l, b) = (46^\circ.3, 53^\circ)$.

SZ-free CMB map delivered in the 2018 legacy release.

Goodies 3) On the SMICA likelihood for CMB

How bad is it to use a stationary Gaussian likelihood in SMICA?

It can only hurt estimation variance, does not introduce bias.

We can evaluate it in the limit of the simplest case: N channels, foregrounds confined to an $N-1$ dimensional subspace, no beam, no noise, perfect calibration.

Then, the SMICA likelihood is 'perfect': the MLE of the best channel combination for extracting the CMB is strictly independent of the model of the foreground distribution!

The only thing that matters here is proper modelling of the CMB distribution which SMICA does well, the CMB being Gaussian stationary.

Summary. SMICA = SM + ICA

- ICA : Independent Component Analysis

. - a.k.a. BSS: Blind Source Separation
. - Component separation using **only** statistical independence
. - Complementary to model-based (parametric) approaches
. - Foregrounds modeled confined to a subspace but otherwise arbitrary.
. - Intermediate between no model at all (ILC, template fitting) and physical/parametric models (Commander).

- SM : Spectral Matching

. - Do it using maximum-likelihood fitting of all auto- and cross-spectra
. - Hence, SMICA does:
. **first** separation of component spectra,
. **then**, and optionally, separation of maps.

Thanks!